

学校编码: 10384

分类号_____密级_____

学号: X2010230501

UDC _____

厦门大学

工 程 硕 士 学 位 论 文

基于随机数算法的自动组卷系统的
设计与实现

Design and Implementation of Auto Generating Test Paper
System Based on Algorithm of Random Number

徐国琴

指 导 教 师: 林 坤 辉 教 授

专 业 名 称: 软 件 工 程

论文提交日期: 2013 年 4 月

论文答辩日期: 2013 年 5 月

学位授予日期: 年 月

指 导 教 师: _____

答辩委员会主席: _____

2013 年 4 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（ ） 1. 经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

（ ☒ ） 2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

摘 要

随着计算机技术及人工智能的发展，组卷系统的研究逐渐被越来越多的专家学者所关注。自动考试系统可减少教师重复劳动，提高工作效率。智能组卷可以提高试卷质量，增加了考试的科学性和客观性。

组卷问题是一个多约束多目标优化问题，它不仅涉及到组卷数学模型建立的问题，还包括对其应用的算法研究。

本文利用.NET 技术和 SQL Server 数据库，设计并实现了基于时间的随机数生成算法的 B/S 结构的自动组卷系统。首先，在对国内外相关文献分析研究的基础上，确定采用 B/S 结构进行研究；其次，在对系统进行功能需求分析的基础上，详细分析了试卷的各项指标，并进一步分析了试题生成过程中试卷的各项约束条件；再次，详细叙述了系统的设计思想、框架结构和程序实现，实现了试题管理、用户管理、自动组卷等功能。最后，研究解决了直接生成 Word 形式试卷。

经过测试和实际运用，本系统运行正常，在减轻教师工作量、提高工作效率等方面具有重要意义，符合教学中自动组卷的实际要求。

关键词：随机数算法； Visual Studio 2008； Microsoft SQL Server 2000

ABSTRACT

As the development of computer technologies and Artificial Intelligence, the research of exam paper auto-generating system has been paid more attentions gradually. The exam paper auto-generating system would be used to relieve teachers of repetitive works, and to improve the work efficiency. The intelligent function of forming exam paper would improve the quality of exam papers and increase the scientific and objectivity of test.

The problem of exam paper generation is a multi-constraint and multi-objective optimization issue. It not only involves proposing the mathematics model of test paper auto-generating, but also includes studying on the application algorithms.

It has designed and implemented the automatic generating test paper system based on B/S structure and algorithm which makes random number based on time, and using .NET and SQL Server database. First, the paper determine to study the system by B/S structure, after analyzing the literature at home and abroad; secondly, we analyzed the role of various indicators of papers and further analysis of the generation process of the paper questions the constraints; thirdly, we describe in detail the system design, framework and processes to achieve, it realizes the functions of examination question management, user management, and automatically making up exam papers.; finally, We have solved the problem of generated the papers in Word.

After testing and practical application, the system is running properly, in reducing the workload of teachers, improving efficiency and so on. It meets requirements of automatic paper in practical teaching.

Key Words: Random Number Algorithm; Visual Studio 2008; Microsoft SQL Server 2000

目 录

第一章 绪论	1
1.1 研究背景与意义	1
1.1.1 研究背景	1
1.1.2 国内外研究状况	2
1.2 本文的工作	4
1.3 论文结构	5
第二章 系统开发的相关技术	7
2.1 C/S 与 B/S 结构	7
2.2 ASP.NET	8
2.3 C#	9
2.4 SQL Server 2000	10
2.5 本章小结	11
第三章 系统的需求分析	12
3.1 组卷的功能需求	12
3.1.1 应用需求与用例图	12
3.1.2 数据流分析	14
3.2 试题属性	15
3.3 实体-关系模型	17
3.4 系统开发方案及运行环境	19
3.5 本章小结	20
第四章 系统的总体设计	21
4.1 系统的体系结构	21
4.2 系统的功能模块介绍	22
4.3 系统的 IPO 图	23
4.4 本章小结	24
第五章 系统的详细设计	25

5.1 系统的功能模块设计	25
5.2 组卷算法的设计	26
5.3 数据库设计	28
5.4 用户界面设计	35
5.5 本章小结	35
第六章 系统实现	36
6.1 软硬件环境	36
6.2 自动组卷算法的实现	36
6.2.1 基于时间的随机数生成算法.....	36
6.2.2 基于随机数的抽题算法.....	38
6.3 功能模块的实现	39
6.3.1 用户登录模块.....	39
6.3.2 试题管理模块.....	41
6.3.3 试题添加模块.....	42
6.3.4 组卷页面.....	43
6.3.5 生成试卷.....	45
6.4 本章小结	47
第七章 系统测试	48
第八章 总结与展望	53
8.1 总结	53
8.2 展望	54
参考文献.....	55
致 谢.....	56

CONTENTS

Chapter 1 Introduction.....	1
1.1 Background of the Research and Significance	1
1.1.1 Background of the Research	1
1.1.2 the Investigation Status	2
1.2 The Work of This Dissertation.....	4
1.3 The Structure of This Dissertation	5
Chapter 2 The Techniques of System Development	7
2.1 C/S and B/S	7
2.2 ASP.NET	8
2.3 C#.....	9
2.4 SQL Server 2000.....	10
2.5 Summary.....	11
Chapter 3 System Requirements Analysis.....	12
3.1 Functional Requirements of System.....	12
3.1.1 Applying Requirement and Use Case Diagram	12
3.1.2 Data Flow Analysis	14
3.2 Property of Exam	15
3.3 Entity-Relationship Model	17
3.4 Develop Concepts and Operating Condition.....	19
3.5 Summary.....	20
Chapter 4 Architecture of System	21
4.1 Architecture	21
4.2 Introduction of Function Modules.....	22
4.3 IPO	23
4.4 Summary.....	24
Chapter 5 System Detailed Design	25
5.1 Design of Function Modules.....	25

5.2 Test Paper Generating Algorithm Design	26
5.3 Database Design of System.....	28
5.4 User-Interface Design	35
5.5 Summary.....	35
Chapter 6 System Implementation.....	36
6.1 Hardware and Software Environment.....	36
6.2 Test Paper Generating Algorithm Realization	36
6.2.1 The Algorithm of Random Number Based on Time	36
6.2.2 Algorithm of Extracting Questions	38
6.3 Function Realization	39
6.3.1 User Login Module	39
6.3.2 Question Management Module.....	41
6.3.3 Add Question Module	42
6.3.4 Test Paper Generating	43
6.3.5 Word Paper.....	45
6.4 Summary.....	47
Chapter 7 System Testing.....	48
Chapter 8 Conclussions and Prospect	53
8.1 Conclussions.....	53
8.2 Prospect.....	54
References	55
Acknowledgments	56

第一章 绪论

随着高校教学改革进一步的深入和大学本科课程建设的逐步完善，对学生掌握每一课程内容程度的考试必须规范化、系统化、科学化和现代化，因此，教学管理也必须现代化和规范化。

1.1 研究背景与意义

1.1.1 研究背景

信息时代来临，促进教育现代化的发展，也促进现代教育的形成。现代化教育不是单指一种行为，步骤或方法上的现代化，而是概括教育的思想，理论，结构，内容，方法，技术，管理等方面的综合范畴。考试是教学过程中的一个重要环节，它既可以检测教学效果和学生对课程的掌握情况，以便在以后的教学中进行必要的调整；又是衡量授课教师教学质量的一种手段。而其命题的好坏又是能否达到预期教学效果的关键因素，因此组卷问题在考试中处于极其重要的位置，是考试过程中的核心环节。以往教学考试中一直采用人工的出题方式。许多学校投入大量人力物力到试题的准备，试题的抽取，试题的组织，试题排版，印刷等过程。这种手工方式出题，缺乏大量试题作参考，教师只能靠个人的经验与知识的积累，另外不同教师所遵循的命题原则不同，评分标准和评分方法不同，命题是带有较大的主观性，不够客观和准确，因此，难以实现教考分离，并且这种命题过程存在着大量人力财力的浪费现象，考试命题质量和工作效率都难以提高，并且不能保证出题的科学性。随着经济，社会的发展及终身教育观念的普及和计算机技术的发展，出现了试题库管理系统。试题库管理系统采用计算机信息管理技术与数据库技术来建设和管理试题库，每次考试前根据任课老师的要求，从题库中选择题目，自动生成试卷。这样排除了在考试过程中人为因素的影像，确保了考试的质量，维护考试的客观性，公正性和准确性，有助于对教学工作的宏观指导和管理，将课程教学的科学管理变成现实。同时，减轻教师的命题工作负担，将广大教师从繁重的教学工作中解脱出来，

更多的投身于研究教学方法和提高教学质量上。目前迅速发展起来的自动组卷系统是计算机科学,教育测量理论和人工智能理论相结合的产物,是从传统教育向现代教育发展的一种新型教育考评体系,是我国教育考试体系的重要组成部分。目前在各级各类大规模教育考试,行业专门技术人员的水平考试,职业技术的水平测试以及聘用干部的测试中占有重要的地位。因此实现自动组卷系统是我国教学考试管理发展的必然趋势。

1.1.2 国内外研究状况

在国外,一些教育相对发达的国家在题库建设方面起步较早,发展也较快,他们对题库理论和模型进行了大量的研究,建立了一些大型题库,20世纪80年代就在一些大规模等级考试中进行了实践应用。典型的应用有美国的TOEFL、研究生入学考试(Graduate Record Examination, GRE)、工商管理类自动组卷系统的设计与实现研究生入学考试(Graduate for Management and Administration Test, GMAT)等。

在国内,题库建设起步比较晚,对其理论的研究直到90年代才受到关注。目前,一些高校也自主研发了一些自己的计算机考试系统,其中比较有代表性的有西安交通大学等联合编制的“高等学校工科大学物理课程试题库系统”、清华大学等联合编制的“高等学校工科高等数学课程试题库系统”、山东省高教自考办公室等联合编制的“高等数学(财经类)题库系统”以及全国计算机等级考试系统等等^[1],但这些系统都存在着不足之处,特别是在组卷算法方面还有很多不完善的地方。

在题库考试系统中组卷模块是其重要的组成部分,而模块中自动组卷功能的实现又是题库考试系统必须实现的核心目标之一,该系统最终所生成的试卷是否高效、科学、合理是评价一个题库考试系统至关重要的标准。

目前在一般考试系统中,组卷模块的算法主要有三种:第一种是将题库看作是试题集,在组卷时由教师根据需要逐题筛选抽取来完成组卷,这种方法虽然也是以计算机作为工具来组卷的,但其仍然没有改变人工组卷的本质,仍然无法摆脱人工组卷所带来的人力资源浪费、速度慢、主观性太强等种种弊端;第二种方法是将题库看作试卷集,需要事先出好多份试卷形成试卷库,在考试

时由组织者从中随机抽取一份或多份试卷进行测试,这种方法由于题库中试卷数量是有限的,在多次考试中容易出现试卷完全相同的情况或者一份试卷只能使用一次,需要不断补充题库,这样造成资源的大量浪费。这两种组卷方法没有实质上的区别,对现在的研究没有任何意义;第三种是有自动抽题组卷的功能,利用某种组卷算法自动形成试卷,利用这种方法生成的试卷灵活性大,但是如果采取的组卷策略不理想,那么生成的试卷就可能不能满足考试的实际要求,在自动组卷后还需要人为的对组卷过程进行干预。因此,为了提高题库考试系统组卷模块的实用性和高效性,组卷策略的选择非常重要。

根据所使用的组卷策略不同目前的组卷系统常见的有以下三种:基于随机函数选取法的自动组卷、基于回溯试探法的自动组卷和基于遗传算法的自动组卷。

随机函数选取法^[2,3]:根据组卷问题的约束条件,由计算机随机地从题库中抽取一道试题,然后判断是否符合试卷的约束条件,不断重复此过程,直到整个组卷工作完成或已经无法再从题库中抽取满足约束条件的试题为止。这种方法操作过程简单,对于单个试题的抽取运行速度较快,但是对于完整的一张试卷的组卷过程来说组卷成功率低,即使组卷成功,所需要花费的时间也是很长的。

回溯试探法^[4]:是建立在随机选取法基础上一种算法。首先它将随机选取法所产生的每一状态类型都记录下来,如果搜索失败就释放上次记录的状态类型,其次再依据一定的规律变换另外一种新的状态类型进行试探,就这样通过不断的回溯试探直到试卷生成完毕或退回出发点为止^[5]。这种算法在回溯试探时所依据的规律是不变的,因此破坏了选取试题的随机性。这种有条件的深度优先算法,在遇到状态类型和出题量都较少的题库系统时,组卷的成功率较好,但在实际应用时选取的试题缺乏随机性,对内存的占用量大,程序结构较复杂且组卷时间也较长。

遗传算法(Genetic Algorithm, GA):是模拟达尔文的生物在自然环境中遗传选择和自然淘汰的进化过程^[6]形成的一种自适应全局优化搜索算法。它的产生源于 20 世纪 60 年代末美国 Michigan 大学的 Holland 在人工适应系统^[7]中设计的一种基于自然演化原理搜索机制。该算法利用一定的编码技术和繁殖机制来

处理复杂的问题，简单通用，适合于并行处理，尤其适用于处理传统搜索方法难于解决的复杂问题和非线性问题，这些显著特点奠定了它作为 21 世纪重要的智能组卷算法之一的地位。在当前的题库自动组卷算法的研究中，许多研究者利用遗传算法在自动组卷方法的应用方面展开了大量的研究^[8, 9, 10]。遗传算法在上个世纪 80 年代进入了蓬勃发展时期，人们在理论和应用方面的研究都十分活跃，并且被广泛应用于各种复杂系统的自适应控制及复杂的优化问题中，取得了良好的效果。美国于 1985 年召开了第一届国际遗传算法会议(简称“ICGA”)，并成立了国际 GA 学会。Goldberg 在 1989 年出版的《Genetic Algorithms in Search, optimization, and Machine Learning》中总结了 GA 研究的主要成果，对 GA 及其应用作了全面而系统的论述，奠定了现代 GA 的基础。L. Davis 于 1991 年编辑出版的《Handbook of Genetic Algorithms》一书中阐述了 GA 在工程技术和社会生活中的大量应用实例，该书对遗传算法的有效应用具有重要指导作用^[11]。近年来，我国有关遗传算法的应用在许多领域取得了令人瞩目的成果^[12, 13, 14]，并且在算法改进及理论研究^[15]方面也作出了成功的探索。

从技术上讲，目前题库系统大体有三种设计模式：C/S 结构模式、B/S 结构模式以及 C/S 结构模式和 B/S 结构模式的混合编程系统。这三种模式互有利弊：C/S 结构模式的题库系统安全性不能得到保证，C/S 结构模式和 B/S 结构模式的混合编程系统由于目前技术尚不完善，其稳定性不能得到保证，基于这两点考虑，本文采用了 B/S 模式的题库系统。

1.2 本文的工作

目前，在新疆的各学校绝大多数教师出卷仍然是根据教学大纲采用直接在 WORD 中排版的方式。这样不仅试题的重用性差而且难以进行科学的分类和管理，教师出卷也存在着效率低、重复劳动和出现差错的可能性较大等问题。

精品课程建设是各高校都重点投入的项目，对于自动组卷系统更是精品课程建设中的重要组成部分。国内各重点高校均有自己的各类课程的题库，在近期该领域处于较稳定的状态；从长远来看，各种新型计算机的问世是必然的，各种新的结构、新的技术将层出不穷，试题的内容将发生重大变化，自动组卷系统也自然要发生重大变化；当然变化与原有内容具有密不可分的联系，要形

成这些变化与发展在某种程度上依赖于对现在知识的掌握、运用和整理总结。

鉴于此,在参考大量相关资料的基础上,本文设计并实现了“数据结构”精品课程自动组卷系统,以更加合理的考核方式来促进教学质量的提高。

本文通过对组卷系统的发展与研究现状进行系统分析,并利用基于随机数算法开发了一个比较实用的自动组卷系统。主要工作如下:

1. 查阅国内外相关文献,了解国内外各类题库特点,分析系统的功能需求,确定本文所开发系统的设计方案,以及实现各功能的具体步骤,合理地完成整个过程。

2. 深入研究组卷的有关理论、评价体系以及基于随机数算法,实现基于随机数算法的组卷策略进行数学建模,并在此基础上,实现高效、合理、准确地实施组卷策略。

3. 确定本试题库系统采用 ASP.NET 和 Microsoft SQL Server 2000 开发,系统结构为 B/S 结构。

1.3 论文结构

本论文共分为八部分。

第一章 绪论:介绍论文的研究背景、国内外研究现状、本文的主要工作和组织结构等内容。

第二章 系统开发的相关技术:包括 C/S 与 B/S 结构、面向对象技术、ASP.NET 开发技术、Microsoft SQL Server 2000 数据库的选取以及数据库访问技术等相关知识。

第三章 系统的需求分析:从组卷的功能需求出发,建立试题主要属性指标及指标体系,从而确立系统开发方案及运行环境。

第四章 系统的总体结构:阐述自动组卷系统的体系结构并进行小结。

第五章 系统的详细设计:主要从系统设计思路及设计原则、系统的功能模块设计、自动组卷系统的数据库设计、自动组卷系统的用户界面设计等方面进行具体阐述。

第六章 系统的实现与关键技术:首先介绍自动组卷系统的实现的软硬件环境,然后对系统各个功能模块的实现方法以及基于随机数算法的自动组卷功能

实现的关键技术，如随机数算法组卷流程、随机数算法的编码方案设计等方面进行说明。

第七章 系统测试：简单介绍系统实现过程中所采用的测试方法与技术。

第八章 总结与展望：总结全文，并展望后续的研究工作。

厦门大学博硕士论文摘要库

第二章 系统开发的相关技术

近年来，随着计算机技术与网络技术突飞猛进的发展，系统开发也逐渐向网络模式发展。目前主要的系统体系结构有两种模式：客户机/服务器模式（Client/Server，简称 C/S）和 Web 浏览器服务器模式（Browser/Server，简称 B/S）。

2.1 C/S 与 B/S 结构

C/S 是软件系统体系结构，通过它可以充分利用两端硬件环境的优势，将任务合理分配到 Client 端和 Server 端来实现，降低了系统的通讯开销。

B/S 结构即浏览器和服务器结构。它是随着 Internet 技术的兴起，对 C/S 结构的一种变化或者改进的结构。在这种结构下，用户工作界面是通过 WWW 浏览器来实现，极少部分事务逻辑在前端（Browser）实现，但是主要事务逻辑在服务器端实现，形成所谓三层结构。

以下对 C/S 与 B/S 进行比较：

（1）数据安全性比较。由于 C/S 结构软件的数据分布特性，客户端所发生的火灾、盗抢、地震、病毒、黑客等都成了可怕的数据杀手。另外，对于集团级的异地软件应用，C/S 结构的软件必须在各地安装多个服务器，并在多个服务器之间进行数据同步。如此一来，每个数据点上的数据安全都影响了整个应用的数据安全。所以，对于集团级的大型应用来讲，C/S 结构软件的安全性是令人无法接受的。对于 B/S 结构的软件来讲，由于其数据集中存放于总部的数据库服务器，客户端不保存任何业务数据和数据库连接信息，也无需进行什么数据同步，所以这些安全问题也就自然不存在了。

（2）数据一致性比较。在 C/S 结构软件的解决方案里，对于异地经营的大型集团都采用各地安装区域级服务器，然后再进行数据同步的模式。这些服务器每天必须同步完毕之后，总部才可得到最终的数据。由于局部网络故障造成个别数据库不能同步不说，即使同步上来，各服务器也不是一个时点上的数据，数据永远无法一致，不能用于决策。对于 B/S 结构的软件来讲，其数据是集中存放的，客户端发生的每一笔业务单据都直接进入到中央数据库，不存在数据一致性的问

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库